CHRISTIAN ASSOCIATION OF LIBRAIANS IN AFRICA–(KENYA)
P.O.BOX 60954-00200, Tel**: +254-020-2720837 / 8**
Nairobi, Kenya

Preservation and conservation in the digital age; what does the future hold in libraries and archives?

Presentation at the 10th Annual CALA Conference held from 25th-29th July 2011 at BTL Christian International Conference Centre-Ruiru, Nairobi

Theme: Library as a knowledge hub

Sub theme: Preservation and conservation of knowledge in the digital regime

By Kisia Oliver Atsango
(Archivist, St. Paul's University Library)

Print to PDF without this message by purchasing novaPDF (http://www.novapdf.com/)

The question of the archive is not a question of the past. It is a question of the future; the question is the future itself, the question of a response, of a promise and of a responsibility for tomorrow. The archive: if we want to know what that will have meant, we will only know in times to come. Perhaps

(Jacques Derrida, 1995)

Abstract

Libraries and archives have served as centers of instruction and learning since time immemorial through preservation and conservation of vital knowledge, and through availing channels of access and use for the same to all who need. However, the nature of the creation, transmission and retention of knowledge and information sources is changing fundamentally in this era of rapid technological advancements. As a result, many new issues arise for libraries and archives. The records, the books, the libraries and the archives of yester year are not the same as those of today. This paper gives insight into archives management with a special interest in the areas of preservation and conservation. It also establishes the need for collaboration between libraries and archives as an important strategy to curate information and knowledge in books, records and archival information sources for universal memory by building digital collections of our historical materials as well as developing an infrastructure to support the archiving of born-digital material for scholarly communication. Furthermore, the paper highlights challenges of archiving in the digital age and how they are affecting preservation and conservation measures in libraries and archives as they move from a memory based on physical artifacts, to a hybrid digital and physical environment, and then increasingly shift towards new forms of digital memory.

Print to PDF without this message by purchasing novaPDF (http://www.novapdf.com/)

Definition of terms

ISO 15489-1:2001 defines ***records*** as "information created, received, and maintained as evidence and information by an organization or person, in pursuance of legal obligations or in the transaction of business".

Archives refers to a group (or series) of records determined as having continuing (permanent) value and *preserved and conserved* for legal, fiscal and intrinsic reasons. All archives, both paper and electronic, are unique and irreplaceable and must never be destroyed. The term '*records of continuing value'* is the preferred term to use when describing archives and means the same as '*archival records'*. *Archiving* is a computing term that has no relationship to archival concepts and practices. The term usually refers to the transfer of non-current electronic data from a live database to another database or backup device such as tape, disk, CD etc. In another context, the term *Archiving* has been frequently misused by institutional staff when referring to the process of transferring records off site for later destruction, however, records that are eventually destroyed are not considered archives.

Scholarly communication is an umbrella term used to describe the process of academics, scholars and researchers sharing and publishing their research findings so that they are available to the wider academic community and beyond.

Ephemeral is used to refer to transitory manuscripts, documents, records and archives, i.e. those existing or lasting only briefly.

Print to PDF without this message by purchasing novaPDF (http://www.novapdf.com/)

Table of Contents

Abstract i
Definition of terms ii
Table of Contents iii
1.0 Introduction 1
2.0 Archives Management 1
2.1 The History of Archives 2
3.0 The Digital Age (Digital lives in Libraries and Archives) 3
4.0 The Preservation and Conservation idea 6
4.1 Challenges of Digital Preservation 7
4.2 Issues arising in Digital Preservation for Libraries and Archives 7
5.0 Opportunities for Collaboration 8
6.0 Conclusion 9
Bibliography 10

Print to PDF without this message by purchasing novaPDF (http://www.novapdf.com/)

1.0 Introduction

Craven (2008) argues that "archivists by and large look backward: despite pressures of the modern world – search rooms to be staffed, educational programmes to be arranged, funding bids to be written, targets to be met – the stuff of their world is historical and their approach to life is, in many cases, both reflective and harmonious with this view". However, the digital era has brought to life a new perspective of looking at archives and archivists; a critical look from the external environment and from the future *(rather than from the inside and the past)*, that not only touches on the management, but also the preservation and conservation of books, records and archives. The changing nature of the creation, transmission and retention of these sources of information confirms in us the conviction that the time is right for debate and discussion, and the realization that a new set of questions beg us as Librarians and Archivists for new answers. Questions like; what happens when our books are digitized or created online as evolving, networked works *(as in the case of Google books and the E-books we currently subscribe to)*? What happens when our records are in the form of text messages or embedded in social networking *(as in the case of face book, twitter, yahoo mail and Google mail)*? How do we archive these new messages and media? These changes in the nature of the cultural and scholarly record are rapidly changing the nature of the archive and the future position of the archivist in the digital age.

2.0 Archives Management

Books and records are some of the world's oldest and commonest artifacts. Over time, systems and repositories, such as libraries and archives, have been developed for managing our archive of manuscripts, published books and the records of individuals or institutions. According to International Council on Archives (2008), "these archives constitute the memory of nations and of societies; they shape their identity, and are a cornerstone of the information society". Likewise, by providing evidence of human actions and transactions, these archives support administration and underlie the rights of individuals, organizations and states, and by guaranteeing the rights of citizens to access official information and to knowledge of their history, they are fundamental to democracy, accountability and good governance. *Archives Management* majorly centers on preserving and conserving this information and related sources

Print to PDF without this message by purchasing novaPDF (http://www.novapdf.com/)

to ensure that today's information survives for tomorrow, and to bring history to life for everyone. Archives Management as a practice is born from *Records Management* which revolves around the processes a record is taken through from its time of creation to its disposal largely referred to as *a record's life cycle*. Just like a book, a record is now understood to be a material medium that transmits a message across space and time. But distinct from a book, or as a particular type of book, it is a proximate, persistent representation of some activity. The life cycle concept of records management categorizes records as:

- ***Current Records:*** Those used regularly and frequently in the day-to-day work of an institution, and generally will be referred to and used frequently within a current period of action.
- ***Semi-Current Records:*** These are required for work of the institution, but will be referred to on an infrequent basis. In general they may be referred to at least four times a year, but no more frequently than once a month.
- ***Non-Current Records:*** These are no longer required to be referred to in the work of the institution, but contain records that must be retained in accordance with the institution's retention schedule.
- ***Archival Records:*** These are identified by an institution's Archivist as having a long-term historical, cultural or educational significance. Archival records refers to and includes: books, published and unpublished manuscripts, documents, photographs, films, machine readable magnetic and optic media, maps, plans, and sound recordings, drawings, letters, vouchers, papers or any other thing on which information is recorded or stored through the means of graphic, electronic, mechanical, magnetic, or other means.

2.1 The History of Archives

Paulus (2010) asserts that "the history of the archive is a history of changing communication technologies". These changes have been from orality to literacy (beginning about 6,000 years ago); from scroll to codex (beginning about 2,000 years ago); from manuscript to print (beginning about 500 years ago); and from paper to electronic media (beginning about 100 years ago). Throughout recorded history, archives, libraries, and other cultural repositories have evolved to provide access to and preserve traces of the past for the future. In fact, the current forms and functions of these institutions and our present perceptions of them were only shaped

Print to PDF without this message by purchasing novaPDF (http://www.novapdf.com/)

recently. Posner (1993) indicates that distinctions between 'archival' and 'literary documents', in particular, were not clear in olden times or the middle ages. Jimmerson (2009) further argues that it was the printing press, and the ability to mechanically reproduce texts, that helped divide documents into forms "directed toward a mass audience *(books and journals)*" and others "grounded in personal interactions and organizational transactions *(records and archives)*". Grouping of these disciplines into professions only began to occur in the late 19th and early 20th centuries. For example, the American Library Association was organized in 1876; the American Association of Museums in 1906; and the Society of American Archivists in 1936 (Shepherd, 2009). By the end of the 20th century, libraries collected mostly published books and journals which got used, saved, and then more were produced to be collected, used and saved thus resulting to what was commonly referred to as the *Book Cycle*. On the other hand archives were not any different. They collected and saved private records, which then got used and proliferated with time. The result of continued use, immediate as well as long-term access, was what was commonly referred to as the *Archives Cycle* and its iterations. However, the 21st century creation of material in digital form, dispersed across global networks and uncurated, challenges and seems to have broken the traditional book and archives cycles. It is in the 21st century that the information and knowledge based society coined the term *Digital Age*

3.0 The Digital Age (Digital lives in Libraries and Archives)

The digital age is characterized by the declining cost of computation that has made digital technologies accessible nearly to everyone in all parts of the world, from cities in developed industrial nations to rural villages in developing nations. These new technologies have the potential to fundamentally transform how and what people learn throughout their lives.

As such, the digital age is taking us through a paradigm shift in the way that information is used by society and by individuals. Changes in communication methods and media are impacting the nature of the cultural and scholarly record, thus making scholarly communication methods and media undergo transformation. As society's primary providers of information, libraries and archives became early users of the new digital technology with respect to cataloguing and processing management and later for providing information on their collections to the world wide network community (Kaul, 2008). For many centuries, the book was considered the primary means for transmitting information and knowledge. In the digital age, the book is slowly

Print to PDF without this message by purchasing novaPDF (http://www.novapdf.com/)

but surely being eclipsed by the screen. Unlike in the past when Librarians and Archivists had their focus on managing and preserving physically identifiable sources of information, the present and the future now presents to them what is fondly referred to as *Born Digital* materials or the *Digital Lives of Libraries and Archives.* Born-digital resources can best be described as information items originally created *(born)*, managed *(live)* and used *(preserved and conserved)* in digital form. These include:

- ***Digital books:*** Fondly referred to as ***E-books*** and considered the "in-thing" of our libraries today, they are book-length publications in digital form, consisting of text, images, or both, and produced on, published through, and readable on computers or other electronic devices. The Oxford Dictionary of *English* defines the e-book as "an electronic version of a printed book." However, e-books can and indeed many do exist without any printed equivalent, i.e. those that are born digital. Many authors are opting to publish and sale their books because of the ease with which this can be done. This move can be attributed to the fact that traditional print publishing is today considered limiting, expensive and time consuming.
- ***Digital documents:*** Nearly all documents generated in offices are currently in digital form. Organizations grapple with the whether to maintain them on paper or in digital form. This is a basic, but important decision to make as far preservation and conservation is concerned. For those maintained in digital form, standard formats such as the Portable Document Format (PDF) should be used to retain formatting, while separating the documents from the software that created them. The many efforts to capture and preserve the intellectual output of universities in an IRs are developing expertise in this area.
- ***Harvested Web content:*** While the Internet Archive captures snapshots of the Web, institutions may take it upon themselves to do more focused archiving in a more thorough manner. A national library may archive its nation's Web sites. A university may archive its own domain. Archives might harvest from Web sites related to a particular subject or event. Open-source tools developed by the Internet Archive can be used to crawl and provide access to the content. The data can be kept in the ISO standard WARC (WebARChive) file format.
- ***Digital photographs:*** The prevalence of digital cameras is making digital photos one of the fastest growing forms of born-digital content. Custodial emphasis is usually on

Print to PDF without this message by purchasing novaPDF (http://www.novapdf.com/)

ensuring they are in current, mainstream formats and are copied onto contemporary, durable media. Care needs to be taken concerning color space and compression, which may affect the integrity of the photographs. Much is known about digital photographs and there have been years of experience in curating collections of this nature.

- ***Electronic records:*** This category includes government documents and corporate, institutional, and organizational archives. This type of collection might consist mostly of documents in word processing formats or may include an array of e-mail, databases, spreadsheets, presentations, and other types of files, some of which can only be read using proprietary software. In most cases it's best to get the content out of proprietary formats. Archivists should be involved in setting policy for their institutions and not just doing clean-up.
- ***Static data sets:*** Data sets are created in the course of research and can be the basis for future research, but they are often created without consideration for preservation or future access. Some data sets need special software and documentation to make them usable and the system may need to be retained or emulated. Context, including the nature of the sample, data collection approach, and software used, should be retained.
- ***Dynamic data:*** This type includes data sets that are added to over time, time-based, or that include genetic sequencing or computer-aided design (CAD). It can include data that is meaningless until it is acted upon—and there may be an infinite number of actions and results. In many cases the software, if not the hardware, environment will need to be retained or emulated. Dynamic data can also include social environments such as Face book and Twitter.
- ***Digital art:*** Digital art may be as simple as digital photography or it may be much more complex in that it could be mixed media, dynamic, or could require recreation of an entire installation to render it effective.
- ***Digital media publications:*** These are materials that are routinely published in digital form. Commercial publications like music CDs, movies on DVD, and video games are on fairly stable media and when those media are replaced, the content is often rereleased in new formats. Libraries tend to keep up with the formats that their users want. There is little immediate concern here and licensing and copyright make it difficult for libraries or archives to take action. But, as with early motion pictures, at some point the content will

Print to PDF without this message by purchasing novaPDF (http://www.novapdf.com/)

lose its commercial value and, unless someone takes custodial responsibility, it will be lost through obsolescence or decay.

Besides preserving and providing access to born-digital material a great number of archives and libraries nowadays have also turned to creating digital surrogates from their existing resources (Kaul, 2008). This trend towards the future calls for an urgent consideration of the conventional preservation and conservation idea.

4.0 The Preservation and Conservation idea

Preservation is branch of Library and Information Science concerned with maintaining or restoring access to books, artifacts, documents and records through the study, diagnosis, treatment and prevention of decay and damage to minimize the loss of information and to extend the life of cultural property[1]. According to ISO 15489-1:2001, preservation entails processes and operations involved in ensuring the technical and intellectual survival of authentic records through time. In a library setting, selection often implies preservation, but this doesn't automatically guarantee survival. Survival often depends on factors such as physical form, number of copies, popularity, and where books rest *(e.g. on library bookshelves).*

On the other hand, ***Conservation*** refers to the actual treatment and repair of individual items to slow decay or to restore them to usable state[2]. The key areas of the conventional preservation and conservation idea included:

- Identifying and acquiring the original documents
- Providing conservation treatment to deal with chemical and physical problems
- Transferring the information to an alternative medium
- Providing a level of environmental control, housing, care and maintenance that will retard chemical deterioration and protect materials from physical damage

However, the digital age presents us with a detour from what was conventional. The computer has now become the writing environment of choice for many writers thus making preserving and conserving take on new meaning. In this age, preserving and conserving means finding a way to preserve not only computer files, but also the software used to create the files and the operating

[1] A Glossary of Archival and Records Terminology. Society of American Archivists Retrieved June 30, 2011
[2] A Glossary of Archival and Records Terminology. Society of American Archivists Retrieved June 30, 2011

Print to PDF without this message by purchasing novaPDF (http://www.novapdf.com/)

system necessary to run that software. Ceeney (2011) asserts that "digital information is inherently far more ephemeral than paper thus the bigger need to put in place digital preservation measures that make sure that information remains accessible for as long as it is needed, even after the format in which it was created has become obsolete". In the light of this argument, preservation in the digital age therefore presents new challenges for archivists and librarians.

4.1 Challenges of Digital Preservation

The paper system worked very well because it followed the records life cycle concept. Just like a living organism is born, then lives and finally has to die, the life cycle work allowed an easy transition of information sources into archives. However, the electronic work flow is very different because of the existence of multiple versions, multiple owners and multiple locations of what might be critical information. This is already causing headaches in providing evidence-based policy advice and providing audit trails, let alone the challenge of providing historical records. Consequently, most public and private sector information is now created in digital format and much of it is put on the web but, unlike paper records which have life spans of centuries, electronic records are far more vulnerable. The whole nature of information has fundamentally changed. Archivists and librarians are now only really starting to grapple with the issues of how to record and safeguard the information being created.

4.2 Issues arising in Digital Preservation for Libraries and Archives

Although the challenges presented may seem relatively few, a myriad of issues arise from these challenges for librarians and archivists in handling preservation of digital sources:

- First, there still exists the need for *physical preservation* even for digital media which entails maintaining the integrity of the bits that reside on a storage medium considering that no matter what media information is stored on; degradation at one time must set in. This in the digital age is called 'bit rot'. There is also the possibility of losing the media altogether or that of things being easily modified, sometimes very convincingly that the original nature is totally lost.
- Second, there is the need for *logical preservation.* This entails maintaining the integrity of the logical ordering of the bits to make them "renderable" or readable in the future.

Print to PDF without this message by purchasing novaPDF (http://www.novapdf.com/)

This ordering depends on software and hardware, which become obsolete and unsupported with time. A good example is that of the floppy disk drive that was a onetime wonder of storage for computer files. Today, the use of the any kind of diskette is unheard of since no modern day PC comes with a floppy disk drive and therefore, information stored on floppy disks can often no longer be read!

- There is also the more basic question of *responsibility*: who will save what, when, how, and where? Common computer applications and uses do not do much to support long-term access; therefore the threat of digital content becoming ephemeral if it is not proactively curated is compounded.

The future of archives rests squarely on how librarians and archivists can collaboratively combat these challenges and address arising issues. Librarians and archivist have for a long time worked together, and as argued earlier on in this paper, there was at one time no clear distinction between these two professionals. The digital age therefore presents us with a new platform on which we can bring our minds together to exploit arising opportunities.

5.0 Opportunities for Collaboration

The change that the digital age has driven is not just a story of risks, but also a story of major opportunities. The past decade has witnessed phenomenal growth in the relevance and accessibility of archival material. Before the web, people didn't come to a library or an archive unless they knew how to research, a situation which has completely changed. These technologies have made both of these institutions a favorite of not only researchers, academicians and students but also and the general public. The barriers to use have come down and within a click of a button one can download and read the information any information needed, as well as do the many things they couldn't in a library or archives. Changes in scholarly communication are causing libraries to broaden their collecting and distribution activities. Libraries are becoming interested in opportunities to collect new types of digital material, including unpublished faculty works, research data, administrative records, instructional materials and software even from social environments. The Library of Congress has pioneered in this by taking on the Twitter archive. There still is an opportunity for other libraries to archive Face Book and You Tube!! Moving away from a focus on collecting certain types of fixed and final works, they are becoming more interested in the *process* of scholarly communication. Libraries are also involved

Print to PDF without this message by purchasing novaPDF (http://www.novapdf.com/)

in developing new models of scholarly communication and publishing services that support authors and editors. Much confusion and uncertainty prevail today, but issues surrounding the complexity and fragility of digital materials are causing archivists, librarians, and others to focus collaboratively on digital curation. This requires them to become active earlier on in the life cycle of information and it is leading toward a convergence between the missions of libraries and archives. All of this means that libraries and archives are repositioning themselves within the Archival Cycle, closer to the point of creation, where they are able to influence, guide, or control the management of digital materials to ensure better access to the scholarly and cultural record, both immediately, through distribution, and over the long-term, through preservation.

6.0 Conclusion

Anthony Grafton (as quoted in Paulus 2010) pointed out in a *New Yorker* essay that appeared a few years ago, that as libraries become more involved in the creation and dissemination of digital materials, they are returning to an ancient and medieval model of the library that positions libraries and archives to support more broadly the organically related activities of collecting, reading, interpreting, creating, disseminating, and preserving information. Librarians and archivists in the developed world are waking up and acting fast to the reality of the fact that the digital age materials are at much greater risk of either being lost and no longer available as historical resources, or of being altered, preventing future researchers from studying them in their original form. Unfortunately for us in the developing world, there exists an acute lack of awareness that leaves us in a precarious position as far as educating and informing our current and future generations is concerned. We are at the door step of a process that will determine the future of our information, with the choice to adopt and follow suit or to let our future perish.

"The future is not something we enter. The future is something we create".
(Leonard I. Sweet).

Print to PDF without this message by purchasing novaPDF (http://www.novapdf.com/)

Bibliography

119894256812.pdf. (n.d.). . Retrieved from http://www.archives.org.il/UserFiles/File/119894256812.pdf

Anderson, I. G., & Tedd, L. A. (Eds.). (2005). *Digital Histories*. Bradford, GBR: Emerald Group Publishing Ltd. Retrieved from http://site.ebrary.com/lib/stpaulslimuru/docDetail.action?docID=10103484

Cox, R. J., & Wallace, D. A. (Eds.). (2002). *Archives and the Public Good: Accountability and Records in Modern Society*. Westport, CT, USA: Greenwood Press. Retrieved from http://site.ebrary.com/lib/stpaulslimuru/docDetail.action?docID=10023340

Craven, L. (2008). *What Are Archives?: Cultural and Theoretical Perspectives: A Reader*. Abingdon, Oxon, , GBR: Ashgate Publishing Group. Retrieved from http://site.ebrary.com/lib/stpaulslimuru/docDetail.action?docID=10257439

Ernst Posner, *Archives in the Ancient World* (Chicago, Ill.: The Society of American Archivists, 2003), xxxi, 14, 27-28; Lionel Casson, *Libraries in the Ancient World* (New Haven, Conn.: Yale University Press, 2001), 2-3

Jimeerson, R. C. (2009). *Archives Power: Memory, Accountability, and Social Justice*. Chicago: Society of American Archivists.

Kaul, A. K. (2008). *E-Libraries in 21st Century*. Jaipur, IND: Global Media. Retrieved from http://site.ebrary.com/lib/stpaulslimuru/docDetail.action?docID=10415191

Shepherd, E. (2009). *Archives and Archivists in 20th Century England*. Farnham, Surrey, GBR: Ashgate Publishing Group. Retrieved from http://site.ebrary.com/lib/stpaulslimuru/docDetail.action?docID=10325944

Print to PDF without this message by purchasing novaPDF (http://www.novapdf.com/)